A new look at Neat Cement Slurry Properties in the Digital Age: Experimental Investigations and Data Analysis using a Python Package

Nachiket Arbad[1], Opeyemi Bello[1], Fernando Aaron Rincon Silias[1], Catalin Teodoriu[1], Mahmood Amani[2]

[1]Mewbourne School of Petroleum and Geological Engineering, University of Oklahoma, USA

[2]Texas A&M University at Qatar

Keywords

Cement operations, oil, gas and geothermal wells, strength of neat cement, well integrity, Python tool

ABSTRACT

Cementing operations are possibly the most sensitive and important stages in the drilling and completion operations of oil, gas and geothermal wells. An inadequate cement slurry design and the stresses a cement sheath encounters most especially when exploiting High-Temperature and High-Pressure wells can lead to loss of zonal isolation and other related downhole problems, jeopardizing the future of the well. A good cement blend quality and cementing job practices play a significant role in oil, gas and geothermal cementing, with integrity resilient to withstand several temperatures and pressure cycles over the life of the well.

The motivation of this study is to evaluate the remaining useful neat cement strength in an elevated temperature condition as a function of time. An experimental study has been conducted by performing a destructive and a non-destructive test on multiple neat cement samples in the laboratory. Furthermore, the paper investigates and develops correlations using the experimental dataset that will aid in predicting the strength of neat cement undergoing different failure modes. A simple Python-based tool (Scikit-Learn) was used for the observational data analysis using both classification and regression features. The findings indicated that the cement mechanical properties show a better consistency at a higher temperature as compared to the results obtained at a room temperature. Low temperature curing conditions show a wider spread of the unconfined compressive strength values, and thus, this will imply that shallower wells, may be difficult to be evaluated using a reduce number of samples. This may imply that use of statistical approaches to estimate well integrity may differ as a function of curing temperature.

1. Introduction

Oil, gas and geothermal wells are commonly cased and cemented with the most commonly used API types of cement, Classes G and H cements. Wellbore cement plays a major role in providing cost effective and long-term wellbore integrity. A good cementing operation entails protection of casing from corrosive fluids, provide mechanical support and segregate productive zones by ensuring long-term operational wellbore integrity (Didier et al., 2018). The application of Portland cement when subjected to an elevated well temperature (as found in geothermal wells), poses serious well integrity issues. Geothermal wells contain some of the most extreme downhole conditions that a well cement system and added compounds will have to withstand. In practice, a variety of additives are also mixed into the Portland cement. Zhang et al. (2011) cited that Neat cement is mostly of class G well cement, while both pozzolan mix and lightweight cement are the admixtures of cement and fly ash. Lightweight cement is generated using a water/solid ratio that is much greater than commonly used. Iversion et al. (2010) revealed that cement designs for high-temperature geothermal applications have typically 35 to 40% additional crystalline silica to help avert loss of compressive strength and an increase in permeability. Several tests need to be done on the cement in the laboratory prior to using it in the field. One of the important tests is to know the strength of the cement and its compatibility with the drilling mud used. Cement compressive strength (CCS) is one of the properties used to test the reliability and ability of Portland cement to withstand deformation when either subjected to thermal or mechanical load situation. Herianto and Fathaddin (2005) quoted that the compressive strength of a cement system depends on the type of raw materials including additives used, mixture proportions, concrete structure, method and time of curing, and exposure conditions. Furthermore, laboratory experimental techniques have also proven to be the best method to obtain a good slurry design with desired compressive strength during cementing operations.

There have been significant studies in the prediction of cement compressive strength using several techniques in conjunction with laboratory experiments. Carino and Guthrie (1994) provided a critical examination of current standards for testing concrete and to provide the technical basis for their possible modification to improve the reliability of testing high-strength concrete. The experimental study was designed to establish the significance of selected factors on the measured strength of molded, high-strength concrete specimens. A full factorial experiment was designed to examine the effects of cylinder size, end preparation, stress rate and type of testing machine on the measured compressive strength. Machine learning technologies are now employed in many fields to replicate materials behavior. The many prediction techniques proposed so far include empirical models, statistical techniques, and artificial intelligence algorithms (Zain and Abd (2009); Topcu and Saridemir (2009); Tiryaki and Aydin (2014); Yan and Shi (2010)). Akkurt et al. (2003) created a three-layer Genetic Algorithm-Artificial Neural Network (GA-ANN) model for the prediction of 28-day cement strength. Input parameters used in the model creation process included the chemical composition of cement, surface area, particle size distribution, and C_3S and silicate moduli. The developed model was subjected to sensitivity analysis to predict the response of the system to different values of the factors affecting the strength. In a further study, Akkurt et al. (2004) developed a fuzzy logic model to predict the 28-day compressive strength of cement mortar under standard curing conditions. Input parameters used in the model concept procedure included C_3S, SO_3, total alkali, and surface area. The study result revealed that the fuzzy model produced slightly higher error than ANN. Labibzadeh et al. (2010) also studied the effect of pressure and temperature changes in the early compressive strength of oil well class G cement and concluded that faster early-age compressive strength could lead to a reduction in transition phase time (thickening time). Falode et al. (2013) conducted sixteen sets of experiments in the laboratory, and the factorial design method was used to develop a model in predicting the compressive strength of oil well cement and determination of cement slurry behavior in situations where additives are either underused or added in excess of the required quantity. Nowadays, machine learning algorithms (linear/multiple regression, logistic regression, KNN classification, SVM, decision trees, random forest, naïve Bayes' theorem, K-means Clustering, etc.) are found to be more efficient in the development of prediction models compared to traditional statistical methods (Varghese, 2018; EliteDataScience.com). The advantages of using machine learning for generated and huge experimental datasets in the Python environment are attributed to the ease of application, good prediction accuracy, robustness, and analysis of enormous experimental data to gain useful insights.

Currently, a dedicated data base for cement properties is under development at the Well Integrity Laboratory at the University of Oklahoma, and the presented work shows the first interpretation of our data sets.

This study investigates the strength development of both API Classes, C and H cement. For the experimental investigation, a total of 239 cement samples were prepared according to standard industry practice. Uniaxial Compressive Strength (UCS) was measured for all the samples using Test Mark Compressive Strength Testing Machine. A simplified Python-based framework was developed for the obtained experimental data (observational data) to predict the compressive strength of the tested neat cement. Scikit-Learn tool was used for data analysis and visualization because of its accuracy and efficiency in handling enormous datasets. The framework will be adopted for cement health monitoring and prediction trend patterns for cured and tested API well cement samples for geothermal application as a function of time.

2. Materials and Methods

2.1 Material: Sample Preparation

The neat cement used in this experimental study is a paste made with both API class C and H cement samples of 2” × 2” × 2” dimensions. Standard industry practice were followed for the preparation of API Class C cement samples: 600 ml of 14.78 ppg Class C Cement neat recipe consists of 683 g of Class C Cement added to 382.48 g of distilled water in 15 s which was mixed for 35 s using an API Cement mixer. Figure 1 presents the process of sample preparation. With 600 ml of cement recipe, 3 identical cement samples of 2” × 2” × 2” dimensions can be prepared as illustrated in Figure 1. Similarly, 600 ml of 16.36 ppg Class H cement slurry was prepared by mixing 860.26 g of Class H cement to 326.9 g of distilled water. The samples were cured for different durations ranging from few hours to several days at an elevated temperature of 75 °C using a water bath.

2.2 Method: Experimental Set-up

The experiment was conducted based on the API RP 10B-2, recommended Practice for Testing Well Cements. The length and width of the samples was measured using the electronic Vernier

Caliper as presented in Figure 1. Destructive tests to measure the UCS of cement samples were performed using Test Mark Compressive Strength testing machine (see Figure 2) which has an accuracy of ± 0.5%. As demonstrated in Figure 2, the UCS value for each of the tested cement sample is directly displayed on the digital screen. For this experiment, the curing time ranges from a minimum of 4 hours to a maximum of 28 days. 3 samples each for the above-mentioned durations were prepared, as only 3 samples could be prepared from 600 ml of the cement slurry. Since 2” or 50.8 mm molds were used to prepare the cement samples, dimensions of 50.8 ± 0.2 mm were accepted for testing.

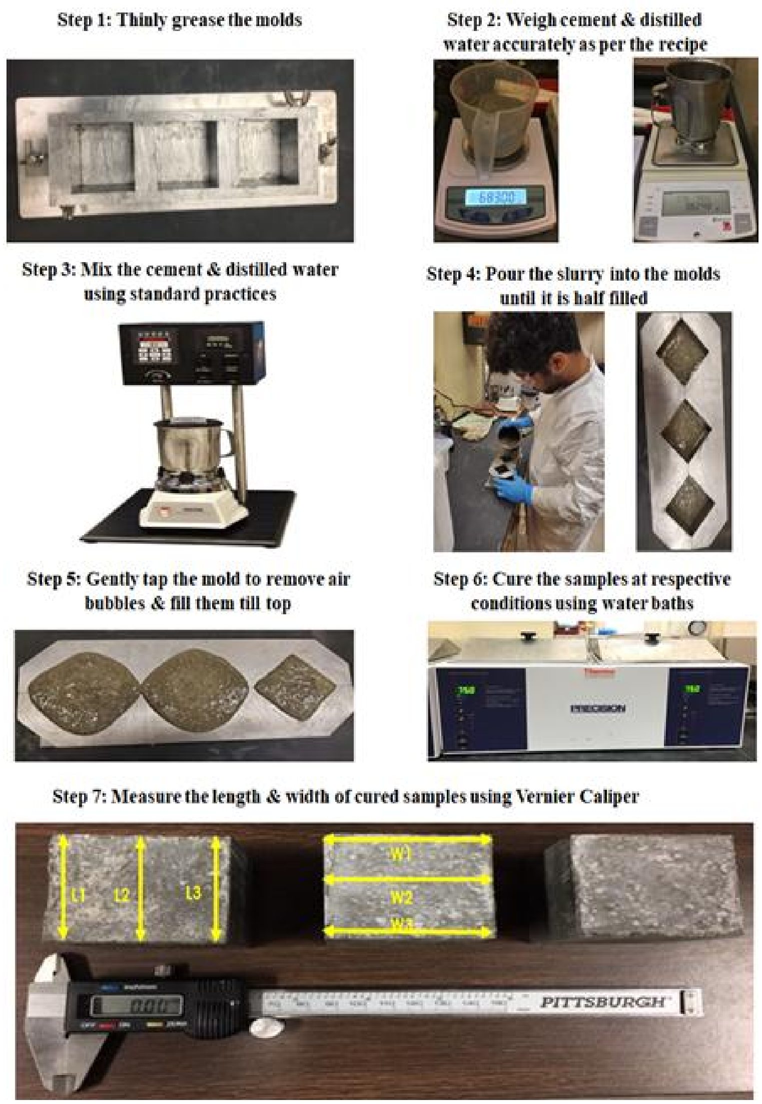

Figure 1: API Class “C” Cement sample preparation procedure

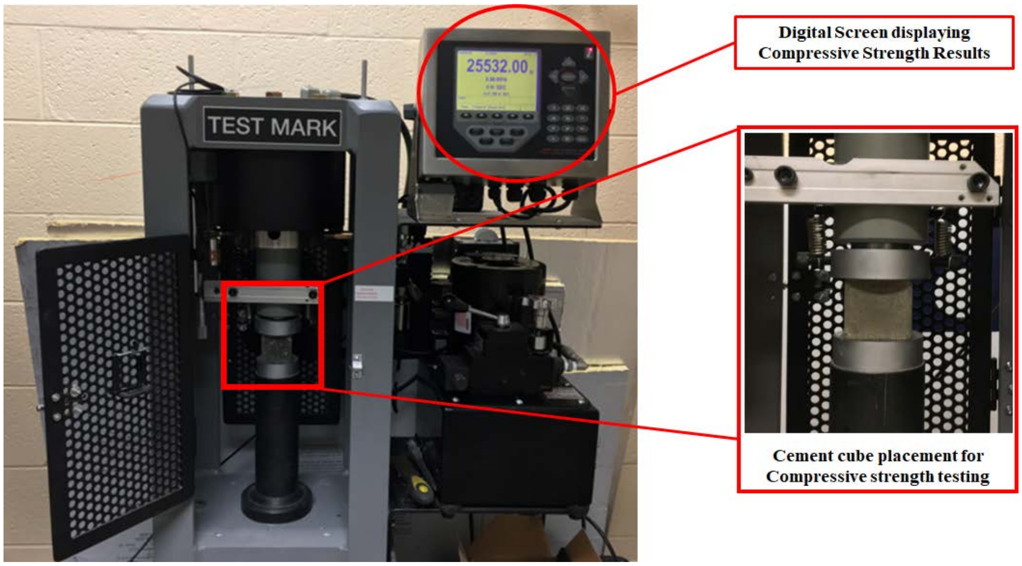

Figure 2: Test Mark Compressive Strength Test Machine

During the experimental test, samples whose dimension were not in the range of 50.8 ± 0.2 mm were prepared again. Dimensional analysis of the cement samples was performed using a descriptive statistics method. Additionally, the length and width of the tested samples were measured thrice for greater accuracy and then the average of these measured three values is used for computation of UCS. Table 1 shows the descriptive statistics of length and width of all samples, while Figure 3 shows Box and Whisker plot as this aids in gaining more information about the outliers with the median being plotted as a dot and the box is delimited by the quartiles and the whiskers by the extreme sample.

Table 1: Descriptive statistics of "Length" and "Width" of the tested samples

Descriptive Statistics	***Length (mm)***	**Width (mm)**
Mean	50.880	50.893
Standard Error	0.008	0.013
Median	50.893	50.883
Mode	50.933	50.873
Standard Deviation	0.091	0.149
Sample Variance	0.008	0.022
Kurtosis	2.442	33.896
Skewness	0.069	4.722
Range	0.633	1.497
Minimum	50.603	50.567
Maximum	51.237	52.063
Sum	6767.043	6768.733
Count	239	239
Confidence Level (95.0%)	0.016	0.026

Figure 3: Box and Whisker plot of Length and Width of the tested samples

2.3 Cement Failure Mode Prediction in Python Environment

The obtained experimental datasets (observational datasets) were employed in Python environment using machine learning library, sklearn with multiple regression methods (MLR) and cluster analysis for the evaluation of the compressive strength for both API Class "C" and Class "H" cements. MLR is one of the statistical methods, which attempts to model the correlation between involving variables and a response variable depending on a linear equation into the observed data. The MLR model is described as:

$$y_i = b_0 + b_{1xi,1} + b_{2xi,2} + \cdots + b_k x_{i,k} + e_i \qquad (1)$$

Where, y_i is the dependent variable; b_0 is the intercept; $x_{i,k}$ is an independent variable; b_k is the vector of regression coefficients; and e_i is random measured errors.

The first step is to import the dataset (experimental data) using **Pandas** as illustrated in Figure 4. **Pandas** is a **Python** open source library for **data science** that allows us to work easily with structured data, such as **csv files**, **SQL tables**, or **Excel spreadsheets**. After importing the csv file, we can print the first five rows of our dataset, the data types of each column as well as the number of null values. **Pandas** provides methods and functions for exploratory data analysis such as, **Dataframe.describe(), Dataframe.info(), Dataframe.dtypes, and Dataframe.shape.**

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import statsmodels.api as sm
import seaborn as sns
sns.set()

from sklearn.linear_model import Linear Regression
```

Load the data

```
data = pd.read_csv('Class C_Cement_view.csv')
```

```
data.head()
```

	Cement Class	System Temperature	Duration	UCS
0	C	75	0.17	8.65
1	C	75	0.17	8.35
2	C	75	0.17	8.40
3	C	75	0.25	11.38
4	C	75	0.25	11.90

Figure 4: sklearn machine learning data frame for API Cement C UCS evaluation

During the supervised learning implementation, our independent variable is Duration (the cement aging), while the UCS is the dependent variable, otherwise known as the target. They are both vectors of length 61 at an elevated temperature of 75 °C (number of observations for cement class C). The same number of vector value was obtained for both cement class C at 25 °C and Class H at 25° and 75 °C, respectively. The cement strength was predicted using single feature. The algorithm identifies the optimal coefficients in our regression model.

3. Results and Discussion

3.1 Linear Regression sklearn Machine Learning Results

The regression table is displayed in Figure 5. The sklearn StatsModel can be divided into two parts namely, "***model summary***" and "***coefficient table***." The model summary comprises of following parameters, dependable variable, the model type used, No of observations, R-squared, Adj. R-squared , F-statistics and Prob(F-statistics). Ordinary least squares (OLS) is the model used for the dependable variable. The coefficient table entails the constant value of the intercept (also known as bias), standard error (the lower the standard error the better the estimate), t-statistics and P-value.

From the result, we can conclude that our model is very significant. Our predicted values for both t-statistics and P-value show a good result. According to the hypothesis, a p-value <0.05 means that the variable is significant, thus our model shows a good fit with the result as 0.000, lower than the common alpha level of 0.05. The zero output implies that the variable, Duration is a significant variable when predicting cement UCS in an elevated temperature environment. This feature strengthens the explanatory power of the computed model. Additionally, our F-statistics value shows an acceptable prediction for the overall testing significance of the model. Using the polyfit feature, the best results were obtained for logarithmic trendline for all the tested samples as presented in Table 2.

Table 2: Summary of Regression Analysis

Cement Class	R-Squared	Adj. R-squared	Std. Error	P-Value
Class C @25°C	**0.975**	**0.975**	**0.187**	**1.26 E-22**
Class C @75°C	**0.807**	**0.804**	**0.633**	**1.68 E-19**
Class H @25°C	**0.965**	**0.964**	**2.22**	**4.9 E-20**
Class H @75°C	**0.895**	**0.894**	**5.61**	**3.0 E-50**

Using the observation data (experimental data), the variation of UCS against the curing time at room temperature and elevated temperature of 75 °C is presented in Figure 6. At elevated temperature for Class "C" cement, the UCS results for more than 14 days were lower than for room temperature, which is somehow unexpected. This could be due to the temperature effect on cement composition, especially for class C. However, for Class "H" cement at elevated temperature, the UCS result is greater as compared to the room temperature condition. The results may indicate a two range for cement behavior: short time (typically less than 14 days) and long term. Currently our team is working on an in-depth evaluation of these two domains.

OLS Regression Results

Dep. Variable:	UCS	R-squared:	0.342
Model:	OLS	Adj. R-squared:	0.331
Method:	Least Squares	F-statistic:	30.64
Date:	Wed, 03 Jun 2020	Prob (F-statistic):	7.51e-07
Time:	17:38:03	Log-Likelihood:	-183.90
No. Observations:	61	AIC:	371.8
Df Residuals:	59	BIC:	376.0
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	21.0011	0.772	27.217	0.000	19.457	22.545
Duration	0.4844	0.088	5.536	0.000	0.309	0.660

Omnibus:	11.515	Durbin-Watson:	0.074
Prob(Omnibus):	0.003	Jarque-Bera (JB):	12.285
Skew:	-1.087	Prob(JB):	0.00215
Kurtosis:	3.324	Cond. No.	10.6

Figure 5: Cement C: Linear Regression sklearn Machine Learning results.

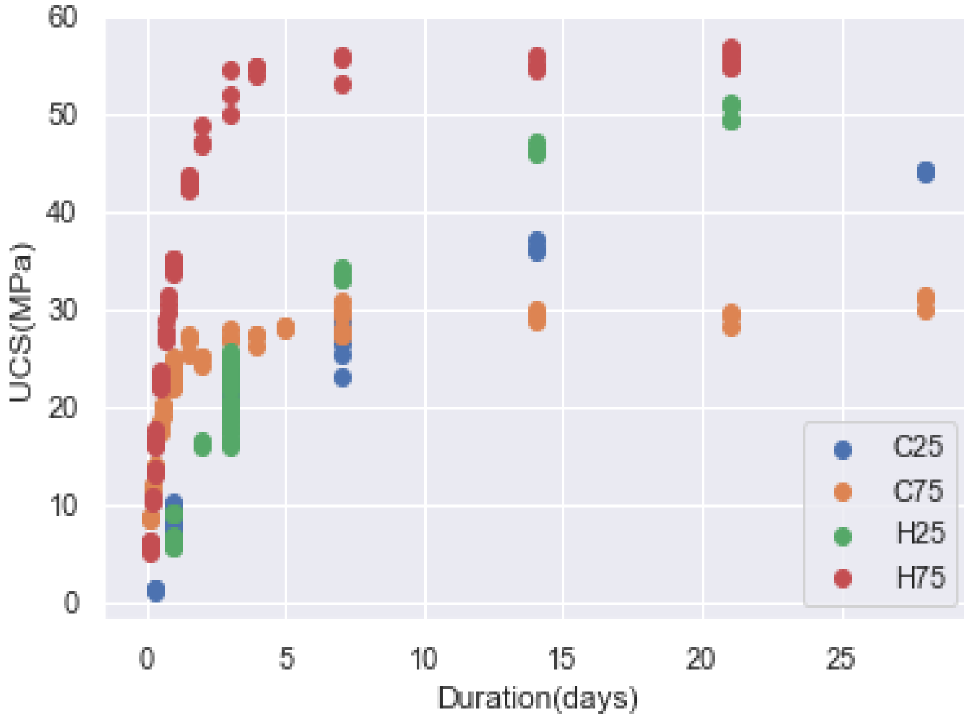

Figure 6: Regression and Classification sklearn Machine Learning results for the tested samples

3.2 Error Analysis

Correlations between UCS and the curing time were obtained from the regression analysis for both Class "C" and Class "H" cements at different curing conditions i.e. room temperature (25 °C) and elevated temperature (75 °C). Error sensitivity analysis was performed by comparing the predicted UCS values with the measured UCS values. Figure 7 shows the error sensitivity analysis results for Class "C" cement samples for room temperature and elevated temperature. The values were overestimated in most of the cases for room temperature as compared to the elevated temperature values. Likewise, Figure 8 shows the results for error sensitivity analysis for Class H samples at room temperature and elevated temperature.

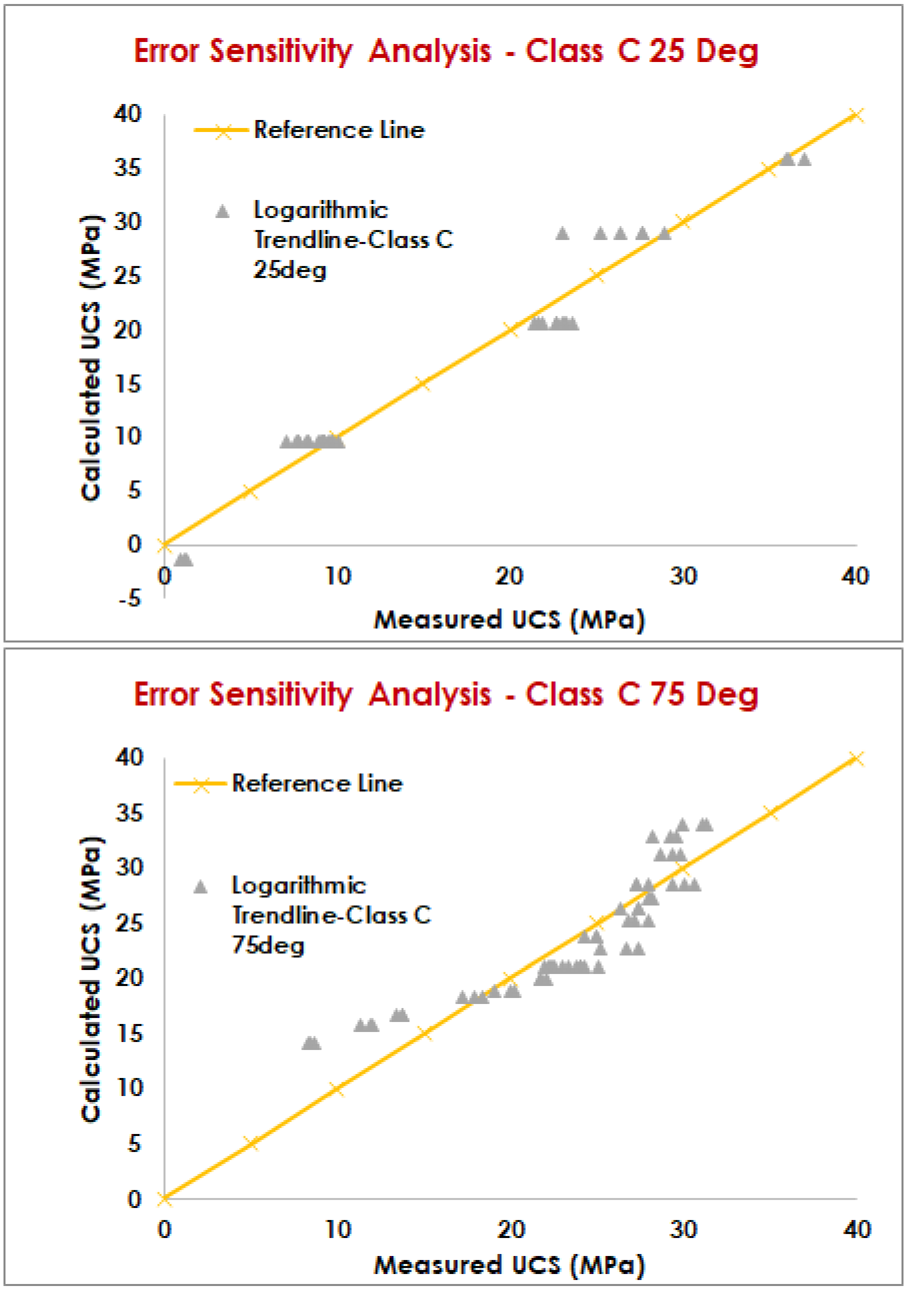

Figure 7: Error Sensitivity Analysis for Class "C" Cement samples

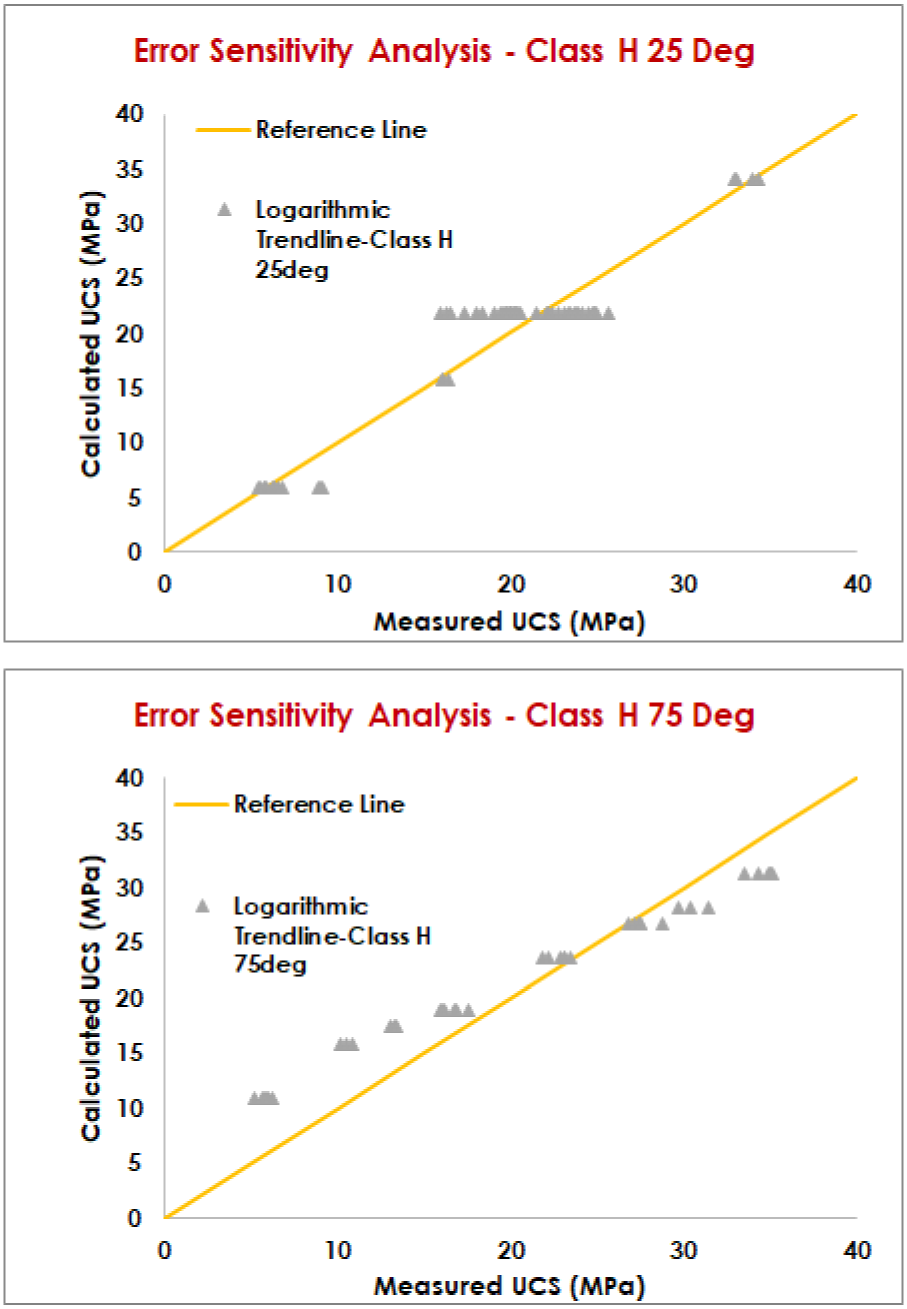

Figure 8: Error Sensitivity Analysis for Class "H" Cement samples

A simple data comparison of 3 Days Class "H" cement UCS revealed that at elevated temperature, the UCS show better consistency as shown in Figure 9, however in depth look at this data shows a higher variation range for room temperature data. Figure 9 shows the UCS results of 60 samples cured for 3 days – 30 samples each at room temperature (25°C) and 75°C respectively. Theoretically the values need to be consistent for same recipe at the same curing conditions i.e. the graph should be horizontal straight line. But practically it is seen that room temperature results have much higher variation then the results obtained for high temperature curing conditions. This variation is not related with the sample geometry variation shown in Figure 3, but rather to the cement batch and probably storing conditions.

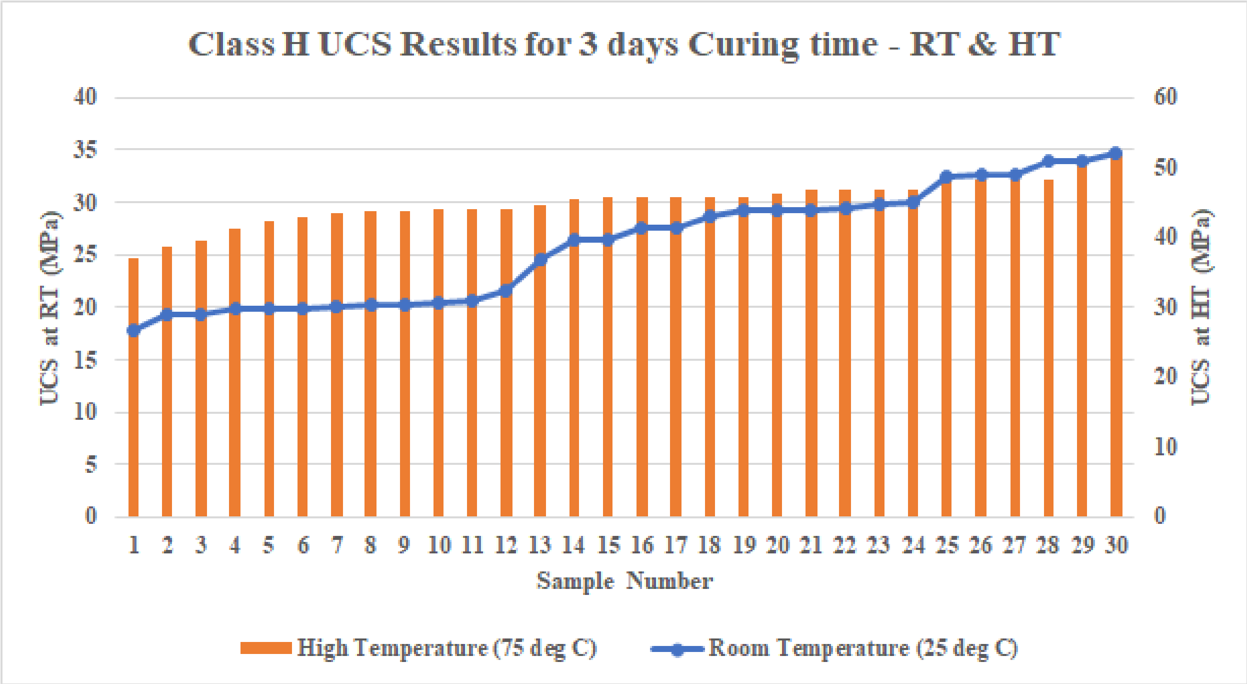

Figure 9: UCS comparison for Class H samples cured at 25°C & 75°C for 3 days curing time.

Conclusions

The following conclusions can be drawn from this study:

i. Laboratory experiments were carried out to obtain data sets for data analysis i.e. UCS values for API Class “C” and “H” cements for different durations at elevated temperature.

ii. The experimental investigation has shown that the UCS measurement and computation are not always depending on the length and width of the samples to be tested, but rather on cement batch and age of the cement powder.

iii. This paper explored the application of a python-based tool for data analysis and visualization to predict cement strength and mechanical behavior. The established framework serves as a template to our future large number of experimental datasets to be developed by focusing on data-driven modeling and health monitoring of tested cement samples under different failure modes. Although the results are promising, more parameters need to be included in future simulations.

iv. Our study findings indicated that the cement mechanical properties show a better error distribution at a higher temperature compared with room conditions. Low temperature curing conditions show a wider spread of the unconfined compressive strength values, and thus, this will imply that shallower wells, may be difficult to be evaluated using a reduced number of samples.

Abbreviations

CO_2	Carbon dioxide
ANN	Artificial Neural Networks
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
API	American Petroleum Institute
UCS	Unconfined Compressive Strength

Acknowledgement

This research work was made in part possible by a National Priorities Research Program grant from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the author.

Many thanks to the Well Integrity Laboratory at OU for granting access to the cement data base.

References

Akkurt, S., Ozdemir, S., Tayfur, G., Akyol, B. "The Use Of GA–ANNs in the Modelling of Compressive Strength of Cement Mortar". *Cement and Concrete Research*, Volume 33, Issue 7, 2003, Pages 973-979. https://doi.org/10.1016/S0008-8846(03)00006-1.

Akkurt, S., Tayfur, G., Can, S. "Fuzzy Logic Model for the Prediction of Cement Compressive Strength". *Cement and Concrete Research*, Volume 34, Issue 8, 2004, Pages 1429-1433. https://doi.org/10.1016/j.cemconres.2004.01.020.

Carino, N.J., Guthrie, W.F. "Effects of Testing Variables on the Strength of High-Strength (90 Mpa) Concrete Cylinders". Book, *U.S. Department of Commerce Technology Administration, National Institute of Standards and Technology*, Gaithersburg, MD 20899, October 1994. https://books.google.com/books?id=ReZIAQAAIAAJ.

Didier, A., Radonjic, M., Du, H. "The Wellbore Cement Additive, Gilsonite, a Solution to Leaky Gas Wells". *Conference Paper: 52nd U.S. Rock Mechanics/Geomechanics Symposium, 17-20 June, Seattle, Washington*. American Rock Mechanics Association, 2018.

Falode, O.A., Salam, K.K., Arinkoola, A.O., Ajagbe, B.M. "Prediction of Compressive Strength of Oil Field Class G Cement Slurry using Factorial Design". *J Petrol Explor Prod Technol 3, 297–302* (2013). https://doi.org/10.1007/s13202-013-0071-0

Iverson, B., Maxson, J., Bour, D. "Strength retrogression in cement under high temperatures conditions". In: *Proceedings of Thirty-Fifth Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, California, February 1-3, 2010, SGP-TR-188 (2010).

Herianto, A., Fathaddin, M.T. "Effects of Additives and Conditioning Time on Comprehensive and Shear Bond Strengths of Geothermal Well Cement. *In: Proceedings of World Geothermal Congress*, Antalya, Turkey, pp 1–7 (2005).

Labibzadeh M, Zahabizadeh B, Khajazfuly A. "Early-Age Compressive Strength Assessment of Oil Well Class G Cement due to Bore Pressure and Temperature Changes. J Am Sci 6(7):38–46 (2010).

Tiryaki, S., Aydın, A. "An Artificial Neural Network Model for Predicting Compression Strength of Heat-Treated Woods and Comparison with a Multiple Linear Regression Model. *Construction and Building Materials, 2014*. 62. 102–108. 10.1016/j.conbuildmat.2014.03.041.

Topçu, I.B., Sarıdemir, M. "Prediction of Compressive Strength of Concrete Containing Fly Ash Using Artificial Neural Networks and Fuzzy Logic". *Computational Materials Science (2009)*. 41. 305-311. 10.1016/j.commatsci.2007.04.009.

Varghese, D. "Comparative Study on Classic Machine Learning Algorithms - Quick Summary on Various ML Algorithms". An Article – Towards Data Science, 2018.

Yan K, Shi C. "Prediction of Elastic Modulus of Normal and High Strength Concrete by Support Vector Machine. Constr Build Mater 2010;24(8):1479–85.

Zain, M., Abd, S.M. "Multiple Regression Model for Compressive Strength Prediction of High-Performance Concrete". *Journal of Applied Sciences*. 9. 10.3923/jas.2009.155.160.

Zhang, M., Faltinson, J.,Palombie, D., Bachu, S. "Review of Regional-Scale In-Situ Conditions and of Well Materials in Alberta". *Final Report submitted to Natural Resources Canada* 2011; 55 p.

https://elitedatascience.com/machine-learning-algorithms